

Re-assessing the role of vowel formant dynamics in speaker-dependent information

An important source of variation in speech originates from the speaker. Earlier research on speaker-dependent information in speech acoustics has investigated speaker information carried by different segments [e.g., 1,2], and by different speech styles [3,4]. This work has for instance shown that vowels tend to carry more speaker-dependent information than consonants. A recent study showed that a single segment within one speech style may vary in speaker-dependent information as a function of the word class it appears in: the vowel /a/ contained more speaker-dependent information when sampled from content than function words [5].

In this study on /a/, however, dynamic formant information did not aid speaker classification, whereas earlier studies of speaker-dependent vowel acoustics named it an important predictor [e.g., 6,7]. In contrast with the earlier work, (1) the /a/ study used spontaneous rather than lab speech, yielding contextual phonetic variation for the vowel /a/, and (2) the vowel /a/ is not inherently dynamic, whereas vowels in earlier work tended to be. Therefore, the present research was aimed at addressing the role of formant dynamics in speaker-dependent vowel acoustics, using a vowel that is often realized inherently dynamically. Also, the study sought to replicate the finding of differential speaker information by word class.

The vowel /e/ (often produced as [eⁱ]) was segmented from spontaneous telephone conversations, spoken by sixty Netherlandish, Standard Dutch, male speakers (~60 tokens/speaker). POS tags and right phonetic context were annotated. Various acoustic measurements were taken, including average and dynamic formant measurements. Preliminary results show that the linguistic-phonetic effects of context and word class are present (assessed through mixed-effects models), and that speaker classification improves when formant dynamics are added (assessed through linear discriminant analysis). This suggests that formant dynamics' contribution to speaker specificity varies by vowel.

References

- [1] Van den Heuvel, H. (1996). *Speaker variability in acoustic properties of Dutch phoneme realisations*. PhD dissertation, Radboud University Nijmegen.
- [2] Andic, A. (2013). *Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning*. PhD dissertation, Radboud University Nijmegen.
- [3] Moos, A. (2010). Long-term formant distributions as a measure of speaker characteristics in read and spontaneous speech. *The Phonetician* 101, 7–24.
- [4] Dellwo, V., Leemann, A., and Kolly, M.-J. (2015). The recognition of read and spontaneous speech in local vernacular: The case of Zurich German. *Journal of Phonetics*, 48, 13–28.
- [5] Heeren, W. (2018). The interaction of linguistic structure and speaker-dependent information in speech. *What is Language conference*, 6–7/12/2018, Zürich, Switzerland.
- [6] McDougall, K. (2006). Dynamic features of speech and the characterization of speakers: towards a new approach using formant frequencies. *International Journal of Speech, Language and the Law* 13(1), 89–126.
- [7] McDougall, K. and Nolan, F. (2007). Discrimination of Speakers Using the Formant Dynamics of /u:/ in British English, In J. Trouvain and W. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, 6–10 August 2007, Saarbrücken, 1825–1828.